molecular systems biology

Synthetic Incoherent Feed-forward Circuits Show Adaptation to the Amount of their Genetic Template.

Leonidas Bleris, Zhen Xie, David Glass, Asa Adadey, Eduardo Sontag, Yaakov Benenson

Corresponding author: Yaakov Benenson, ETH

Review timeline:

Submission date:	08 December 2010
Editorial Decision:	24 January 2011
Revision received:	13 May 2011
Accepted:	06 June 2011

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 24 January 2011

Thank you again for submitting your work to Molecular Systems Biology. We have now heard back from the three referees whom we asked to evaluate your manuscript. As you will see from the reports below, the referees find the topic of your study of potential interest. However, they raise several concerns, which should be convincingly addressed in a revision of the present work.

The main concerns noted by the reviewers refer to the following aspects:
- the need to characterize better the relationship between DNA copy number input and measured fluorescence output.
- additional clarifications are required with regard to the FACS data and the observed behavior of the negative control circuits.

In addition to our capacity to host datasets in our supplementary in formation section, we provide a new functionality on our website, which allows readers to directly download the 'source data' associated with selected figure panels (eg <http://tinyurl.com/365zpej>), for the purpose of alternative visualization, re-analysis or integration with other data. These files are separate from the traditional supplementary information files and are directly linked to specific figure panels. In the case of this study, we would strongly encourage you to submit the source data files corresponding to the quantitative measurements presented in the figures included in this manuscript. We provide below some general guidelines with regard to the format of such data tables.

*** PLEASE NOTE *** As part of the EMBO Publications transparent editorial process initiative (see our Editorial at http://www.nature.com/msb/journal/v6/n1/full/msb201072.html), Molecular Systems Biology will publish online a Review Process File to accompany accepted manuscripts. When preparing your letter of response, please be aware that in the event of acceptance, your cover letter/point-by-point document will be included as part of this File, which will be available to the scientific community. More information about this initiative is available in our Instructions to

Authors. If you have any questions about this initiative, please contact the editorial office msb@embo.org.

If you feel you can satisfactorily deal with these points and those listed by the referees, you may wish to submit a revised version of your manuscript. Please attach a covering letter giving details of the way in which you have handled each of the points raised by the referees.

Thank you for submitting this work. I am looking forward to receiving your revised manuscript.

Yours sincerely,

Editor
Molecular Systems Biology

GENERAL GUIDELINES FOR SUBMISSION OF DATA ASSOCIATED WITH FIGURES:

The data should at least be usable by readers who understand the topic of the paper and should at least allow to redraw (some of) the figures using an alternative graphical representation. The data should also be annotated to some extent to provide essential information about the meaning of the values included. The general purpose of the annotation is to facilitate the link between the data submitted and the description provided in the figure legends in the main paper.

Please note that these are tentative guidelines and we appreciate any feedback you may have.

1. If appropriate standards exist, data and annotation (meta-data) should be provided using the relevant data exchange format and controlled vocabularies, either in XML or using the corresponding tabular format.

2. If no suitable standard is available, data should be submitted in a tabular format. In this case, we recommend to use only a single file per figure panel, containing both the relevant data (time coordinates, measurements, computed values, summary statistics, etc...) and the annotations explaining the meaning of the columns in the data file. The files should be submitted as 'datasets' in the tracking system.

3. If possible, the files should be named according to the associated Figure and panel as follows: 'Source data for Figure 1B,C', 'Source data for Figure 2C and Supplementary Figure S2', etc...

4. Tabular data should be in csv or tab delimited text; data can also be submitted as a single Excel worksheet (1 file per panel), which will be converted into a text file at production stage.

5. The first cell (first row, first column) should contain a brief description (free text, without comma 'tab' or 'return' characters) of the nature of the data and experimental or computational procedure or a reference to the part of the main manuscript which describes the relevant experiment (eg Figure legend, Materials and Methods paragraph).

6. The second row of a table contains column headers (alphanumeric string without space or punctuation); each column has a single data type; empty rows will be deleted.

7. Cells of the third row should contain each a description (free text) of the meaning of the respective columns (do not be afraid to use copy-paste, if necessary); this should specify the *data type* (eg gene symbol, metabolite concentration, phosphorylation level, reporter gene expression, fluorescence level, etc...) and the units (if applicable) of the respective data/measurements. As far as possible, use the same terminology as in figure legend and mention the same labels as on the graphical version of the figure.

8. The columns should be of single data type (eg text, numerical values, dates).

9. There should be no empty rows.

10. The data file should contain, if reasonable, both the raw unprocessed data (corresponding to the

actual experimental measurements or the original computational results, including replicates) and the post-processing values used to generate the figure (eg using summary statistics, averages, standard deviations, etc...).

REFEREE REPORTS

Reviewer #1 (Remarks to the Author):

This is a rigorous study that discovers important and useful circuits that can make biological output insensitive to DNA dosage. The circuits are similar to those commonly found in mammalian cells (as well as in other species) , called incoherent feedforward loops (I1FFL) . In particular, use of RNAi in these feedforward loops makes their robust dosage properties most pronounced. This is relevant since many mammalian I1FFLs seem to use microRNAs in precisely the same positon as in the present paper.

The data is clean and the controls are well done. I liked the use of two fluorescent colors to provide both input and output levels in indidivual cells. I also liked the comparison to feedback loops, which lack the dosage compensation property.

I recommend publication of the paper in its present form.

The authors might like to add references that provide evidence for microRNAs in naturally occuring IFFL circuits, eg:

Tsang J, Zhu J, van Oudenaarden A.
Mol Cell. 2007 Jun 8;26(5):753-67

Reviewer #2 (Remarks to the Author):

Bleris et al constructed sophisticated synthetic circuits to show that incoherent feed-forward loops (IFFLs) enable adaptive gene expression: that is, the level of gene expression adapts to the changes in DNA template abundance. An interesting novelty in these synthetic circuits is the fact that the input level is the DNA itself. The authors implement the IFFLs to work on various levels, transcriptional and post transcriptional, and show that post transcriptional IFFL adapts better with reduced noise levels.
The authors also constructed a negative autoregulatory circuit and provide a detailed computational analysis to compare between the features of the IFFLs and the negative autoregulatory motif.

While the implementation of the various synthetic circuits is impressive, there are several key results that need to be further addressed:

1. A key point of the entire study is the presumed linear relationship between the measured fluorescence and the plasmid copy number. Fluorescence probably positively correlates with plasmid copy number, but is the correlation linear for a wide range of plasmid copy number? The authors cite Schwake et al. to argue that the relationship is linear. But Schwake et al employ a mathematical model to show that the experimental results can be fitted with a linear relationship. Since the linear relationship is central to the interpretation of virtually all the results presented in this study, can the authors supply a better demonstration that in their system fluorescence depends linearly on plasmid number? Moreover, is the liner relationship kept for a wide range of copy number? In particular, is it true for the high levels of copy number where gene expression reaches a steady state and is independent on the plasmid copy number?

On that note, Schwake et al used a similar transfection protocol (Lipofectamin 2000 and similar amount of plasmid). Their conclusion was that the number of active vector complexes (vectors which express a protein) is on the order of one. Can the authors determine the level of active complexes in their system? This is important since the authors assume a wide distribution for the plasmid copy number and all subsequent interpretation of the results as well as the basis for the simulations and noise analysis relies on this assumption. Also, if this is indeed the case, then a significant part of the noise will result from the small number of plasmids. The authors indeed find a significant noise which can not be explained based on their various control circuits.

2. FACS data.
(i) - In fig 3A, the authors show that with full functional repression by LacI (WT), gene expression of the output saturates quickly with respect to the plasmid copy number (input). The adaptive feature gradually decreases when decreasing LacI binding (the different mutants and the negative control). While the increase in the expression of DsRed is obvious when reducing LacI binding, it is not obvious why the maximal expression level of the input (ZsGreen) decreases as the maximal level of DsRed increases. Can the authors explain this observation? One possibility which the authors ought to rule out is that LacI binding to the mutated (and to the negative ctrl) sites of LacO somehow interferes with the expression of the divergent DsRed gene.

(ii) - In Fig 4, the authors show saturation for the post-transcriptional IFFLs. Again, the maximal expression levels for the negative control are significantly lower. In fact, as shown in fig 4c, there is a similar, almost perfect linear relationship, between the input and the output for the IFFL and the negative control for the same range of input levels. Thus, the negative control does not show that it does not adapt as opposed to the 'WT' circuit that reaches saturation. Can the authors explain that? Could it be that if the negative control circuits had higher expression levels a similar saturation will be observed? If so, it will suggest that IFFLs are not required for the adaptive feature in these synthetic circuits and that the adaptation is somehow inherent to this synthetic system.

Also, I could not find any referral to fig 4e in the text. There is also no description of what the two right most images of fig 4e represent (IFFL type I or II ?).

3. FACS analysis
In transformations only a small fraction of the cells eventually obtain the plasmid. Were the non-transformed cells excluded from the analysis? Can the authors show the background level of the cells (non-transformed) for the different channels for the same instrument setup? The authors restricted the analysis to counts which were less than half of the instruments' dynamic range to avoid biasing towards saturated data points, but there is no reference regarding the lower range of the dataset. For example, in figure S2A, it seems that a significant fraction of the cells have considerable DsRed levels (up to 2x10^4 a.u.) but the AmCyan levels are close to zero. Could it be that the lower range of the input-output linear dependence actually falls in the background autofluorescence of the cells?

Minor:
Some of the statements are not clear or don't make sense:
(1) Page 6 (in the middle), the sentence " While strictly speaking...differences in reporter levels" does not sound logic after " synthesis and degradation rates, ". Should either add however or end the sentence and start a new one " Differnces in the copy-number...".

(2) The last sentence starting on page 11 and continues on page 12 does not read well once reading", are slightly sensitive...."

(3) page 13, typo in the 1st paragraph, "outut"
(4) in SI, page 40 typo "explanded"

Reviewer #3 (Remarks to the Author):

The authors analyze synthetic biological circuits in mammalian cells to explore levels of adaptation and fluctuations in output resulting from changes in DNA copy number. They focus on two types of post-transcriptional incoherent feed-forward loops and compare their behavior with a transcriptional FFL and a negative feedback loop. They use numerical and analytical modeling to generate predictions about the behavior of these various circuits. They predicted that all forms of FFLs would show adaptation with respect to DNA copy number variation, but as the inhibition is weakened the input-output behavior would become more linear. The analytical treatment provided additional constraints for when these results would hold true. They tested and validated the predictions experimentally using transient transfection of genetic constructs and flow cytometry. They show that the post-transcriptional IFFLs are less noisy and provide a greater range of adaptability compared with the transcriptional IFFL.

This work presents a clear analysis of a biologically important genetic motif. The manuscript shows a nice, complementary combination of computational, analytical, and experimental techniques to tackle an intriguing question.

Major comments:
1) The authors rely on a fluorescent read-out of the input to the circuit, which is the DNA copy number. However, there is no attempt at calibration between the two quantities. The only mention of this is during the discussion of model in which it is stated, "The amplitude of the measured fluorescence was arbitrarily assigned to correspond to 100 plasmid copies." As far as I know there is no support for such a connection in the literature. Providing some form of calibration would greatly strengthen the manuscript. One possibility would be to use a targeted stable knock-in of the genetic circuit into cells for calibration purposes.
2) The analysis of the "biphasic" regime of the IFFLs is confusing. It seems that for conditions in which such biphasic behavior arises, a higher copy number results in a breakdown of adaptation. However, the experimental analysis of the circuits shows the exact opposite - breakdown of adaptability occurs at low copy number. The authors should address these contradictory results.
3) The results are encouraging for synthetic biologists. However, the conclusion that adaptation occurs only in the larger input regimes reduces the potential impact in regard to gene dosage compensation in natural systems, in which gene copy numbers are very small. Are there other systems which this study can shed light on their behaviors? i.e., systems with greater genetic copy numbers? A discussion of such systems will increase the impact of the manuscript.

Minor comments:
1) It would be helpful to include diagrams and more detailed descriptions of the negative control circuits in the main text, instead of just the Supplementary Materials. It was especially difficult to track down the negative control for the post-transcriptional IFFLs in Figure 4A and B.
2) Why are there different ranges for the inputs in the different experiments in Fig. 3A? One would expect that under the same transfection conditions, expression of the green fluorophore should have a comparable range, since the mutation in the LacO should only affect DsRed expression.
3) I believe the caption of Supp Fig S2N should refer to red and black, rather than green and brown.

1st Revision - authors' response 13 May 2011

Thank you for the invitation to submit a revised version of our manuscript, Synthetic Incoherent Feed-forward Circuits Show Adaptation to the Amount of their Genetic Template. We appreciate reviewers' comments and suggestions. As a result of their feedback we performed a number of additional experiments as well as an additional in-depth study of published literature. Our detailed responses can be found in the point-by-point document. Briefly, we contend that we were able to positively address all the concerns, including the question regarding the linearity of copy number versus fluorescence (that we and others found to be linear) and the issue of reduced fluorescent intensity of the input marker in the control circuits (that we found to likely result from limited ribosome availability). We believe that these new data support our original conclusions and result in a significant improvement over our original submission.

Point-by-point response to reviewers' comments

We appreciate all the comments and remarks and thank the reviewers for an in-depth analysis of our manuscript. We believe that we have been able to address all reviewers' concerns, as detailed in the enclosed point-by-point response and as reflected in the revised version of this manuscript. Our additional analysis corroborated our initial assertions, and the inclusion of the new data has indeed strengthened our conclusions. We hope that the reviewers will find our answers satisfactory. Please note that in the enclosed document the original comments are shown in italics.

For all co-authors,

Yaakov Benenson.

Reviewer #1 (Remarks to the Author):

This is a rigorous study that discovers important and useful circuits that can make biological output insensitive to DNA dosage. The circuits are similar to those commonly found in mammalian cells (as well as in other species), called incoherent feedforward loops (I1FFL). In particular, use of RNAi in these feedforward loops makes their robust dosage properties most pronounced. This is relevant since many mammalian I1FFLs seem to use microRNAs in precisely the same position as in the present paper.

The data is clean and the controls are well done. I liked the use of two fluorescent colors to provide both input and output levels in individual cells. I also liked the comparison to feedback loops, which lack the dosage compensation property.

I recommend publication of the paper in its present form.

Response:

Thank you!

Comment 1.1

The authors might like to add references that provide evidence for microRNAs in naturally occuring IFFL circuits, eg:

Tsang J, Zhu J, van Oudenaarden A. Mol Cell. 2007 Jun 8;26(5):753-67

Response 1.1

We thank the reviewer for his remarks and his recommendation. Indeed the results of this manuscript are relevant to our work and we have added the following references in the paper (new references in bold):

1. An affirmative answer to this conjecture would add experimental support to the hypothesis (Veitia et al, 2008) that similar circuitry might be employed by cells to implement gene dosage compensation in the context of neuronal homeostasis **(Tsang et al, 2007),** sex determination (Lucchesi et al, 2005) and ploidy changes.
2. One example, consistent with our hypothesis, is of an activator that upregulates a gene and a microRNA that then downregulates that same gene (**Nakamoto et al, 2005; Tsang et al, 2007**).
3. While many dosage compensation processes involve genome-wide changes, it has been hypothesized that feedback and negative feed-forward effects could play a role in some cases **(Stenberg et al, 2009; Tsang et al, 2007; Veitia et al, 2008**), and mathematical analysis revealed that a two-component genetic circuit with elements of opposite regulatory activity (activator and inhibitor) constitutes a minimal requirement for network-dosage invariance (Acar et al, 2010).

Reviewer #2 (Remarks to the Author):

Bleris et al constructed sophisticated synthetic circuits to show that incoherent feed-forward loops (IFFLs) enable adaptive gene expression: that is, the level of gene expression adapts to the changes in DNA template abundance. An interesting novelty in these synthetic circuits is the fact that the input level is the DNA itself. The authors implement the IFFLs to work on various levels, transcriptional and post transcriptional, and show that post transcriptional IFFL adapts better with reduced noise levels. The authors also constructed a negative autoregulatory circuit and provide a detailed computational analysis to compare between the features of the IFFLs and the negative autoregulatory motif.

While the implementation of the various synthetic circuits is impressive,

Thank you!

there are several key results that need to be further addressed:

Comment 2.1i

1(i). A key point of the entire study is the presumed linear relationship between the measured fluorescence and the plasmid copy number. Fluorescence probably positively correlates with plasmid copy number, but is the correlation linear for a wide range of plasmid copy number? The authors cite Schwake et al. to argue that the relationship is linear. But Schwake et al employ a mathematical model to show that the experimental results can be fitted with a linear relationship. Since the linear relationship is central to the interpretation of virtually all the results presented in this study, can the authors supply a better

demonstration that in their system fluorescence depends linearly on plasmid number? Moreover, is the liner relationship kept for a wide range of copy number? In particular, is it true for the high levels of copy number where gene expression reaches a steady state and is independent on the plasmid copy number?

Response 2.1i

We present two lines of evidence that support the assumption that expression level varies linearly with the gene copy number. One is based on the analysis of published literature (including some papers we were not aware of during the original submission), and another is a result of additional experimental work we performed.

A. Evidence based on published papers.

The most relevant publication seems to be

Paper 1: Transfection by Cationic Liposomes Using Simultaneous Single Cell Measurements of Plasmid Delivery and Transgene Expression by Wen-Chi Tseng, Frederick. R. Haselton, and Todd D. Giorgio (J. Biol. Chem., 272, pp 25641-25647).

In that work the authors devised a dual-reporter assay in which transfected GFP-expressing plasmids were themselves labeled with a different fluorescent dye. This allowed simultaneous measurement of both plasmid copy number and gene expression in individual cells using flow cytometry. The most relevant data from this paper, summarized in Fig. 8, is reproduced here:

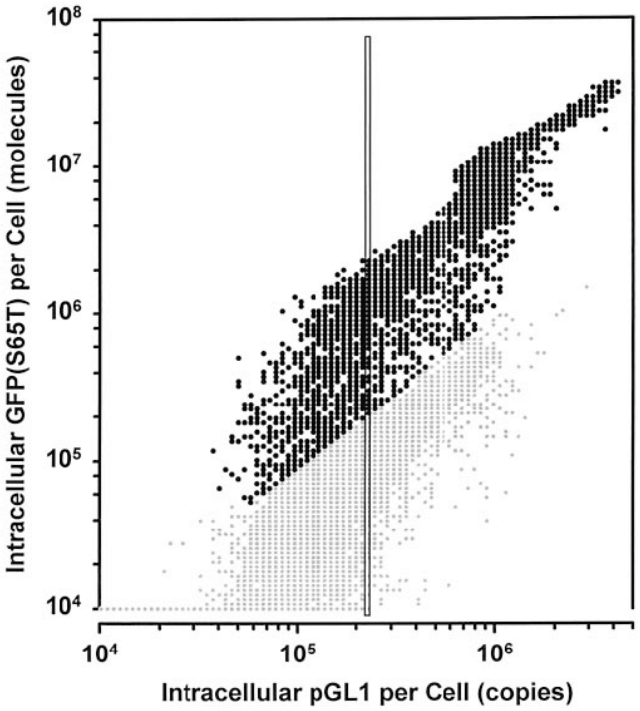
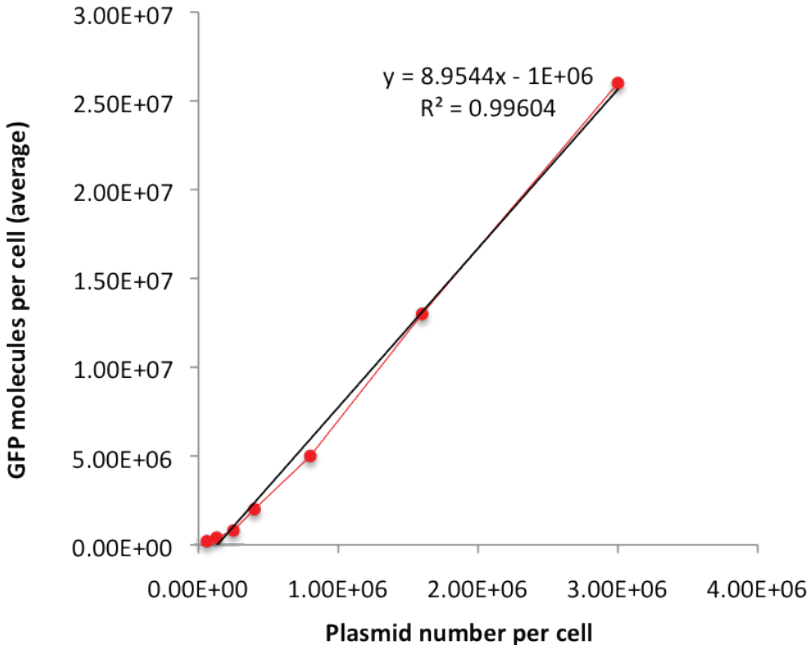
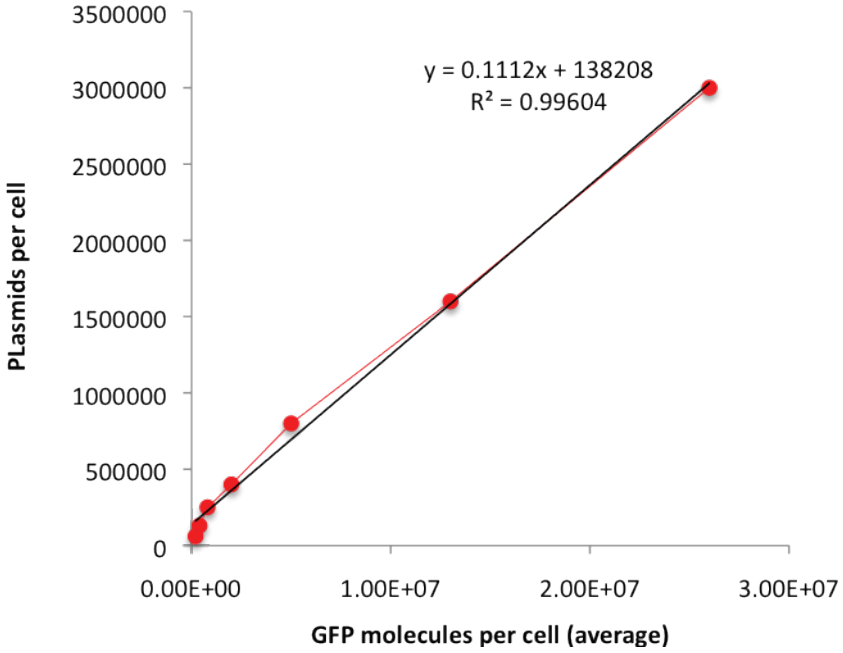


FIG. 8. **Simultaneous measurements of intracellular plasmid copy number and GFP(S65T) synthesis in individual cells.** Intracellular plasmid copy number and GFP(S65T) expression were simultaneously estimated in 20,000 single cells by flow cytometry. Cells were exposed to 1000 ng/2000 ng for 24 h and harvested 24 h post-exposure. Cells which are positive for GFP(S65T) synthesis are identified by *black dots*, while cells whose autofluorescence masks detectable GFP(S65T) are shown as *gray dots*. Cells that express GFP(S65T) strongly also contain a high intracellular pGL1 copy number. Subsequent calculations, shown in Fig. 9, were performed on subpopulations that were one channel wide in plasmid copy number and all 128 channels wide in GFP(S65T) number. One example of such a subpopulation is shown here as a *rectangle*. This is a representative result selected from five similar experiments of identical treatments.

Although the graph uses a log-log scale and the raw data is not available, we manually extracted median GFP levels for different plasmid copy numbers and obtained the following plots (the top plot follows the axis arrangement of the figure in the paper; the bottom plot depicts the number of plasmids as a function of GFP level, which corresponds to the measurement we performed that is described later):





Two features are apparent from these plots. First, GFP levels depend essentially linearly on the copy number. If anything, there is a threshold effect at the low copy number range. Second, the absolute number of copies measured in singe cells is huge – on the order of millions. We analyzed additional manuscripts to address these reported values. It turns out that with lipofection-based methods, the proportion of transcriptionally-active plasmids among the total number of plasmids that reach the cell is extremely low. Major loss of yield occurs in the transition from the cytoplasm to the nucleus. One published report specifically addresses this issue:

Paper 2: Quantification of plasmid DNA copies in the nucleus after lipoplex and polyplex transfection by Richard N. Cohen, Marieke A.E.M. van der Aa, Nichole Macaraeg, Ai Ping Lee, and Francis C. Szoka Jr. (Journal of Controlled Release, 135, pp 166-174 (2009)).

These authors performed careful fractionation of cell content into nuclear and cytoplasmic fraction, with the goal to measure the amount of plasmid delivered into the nucleus. A relevant figure from this paper is reproduced here:

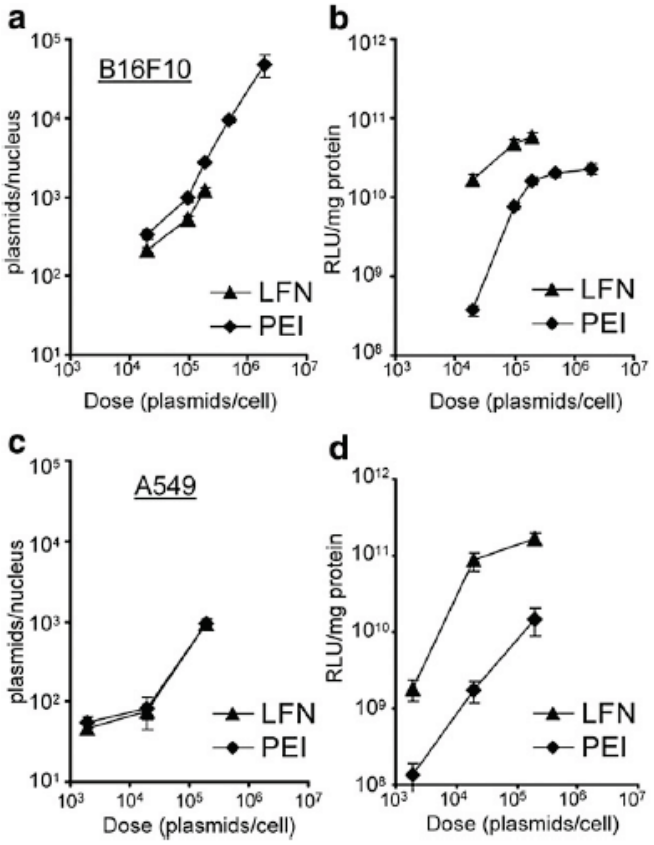


Fig. 4. Plasmids/nucleus and luciferase expression dose response of PEI polyplexes and LFN lipoplexes. B16F10 (a–b) or A549 (c–d) cells were transfected with PEI/pLuc polyplexes or LFN/pLuc lipoplexes at the indicated doses. (a,c) Detected plasmids/nucleus in nuclei isolated by the iodixanol method. (b,d) Detected luciferase expression in lysates from transfected cells using the Promega Luciferase Assay kit normalized to the amount of total protein determined by the Bradford assay. Error bars represent the standard deviation of triplicates in the qPCR assay or luciferase assay.

By visual examination of the axis values in panels **a** and **c**, one observes that nuclear plasmid fraction is less than 1% of the total cellular plasmid content. For simplicity, we can assume that one in a 100 plasmids reaches the nucleus. However, applying 1:100 ratio to the above Fig. 8 would mean that highly-expressing cells have on the order of 30,000 active plasmids in the nucleus, which is clearly a gross overestimation. To estimate how many plasmids are transcriptionally active in the nucleus, we analyzed another report,

Paper 3: Polyethylenimine but Not Cationic Lipids Promotes Transgene Delivery to the Nucleus in Mammalian Cells by H. Pollard, J.-S. Remy, G. Loussouarn, S. Demolombe, J.-P. Behr, and D. Escande (J. Biol. Chem., 273, pp 7505-7511, 1998)

Here the authors performed direct injection of naked cDNA into nuclei of mammalian cells. Their cDNA expressed b-galactosidase and gene expression was measured in On/Off terms by X-gal staining. While the original goal of the paper was to compare gene expression from naked and complexed DNA, one piece of evidence hints on the efficiency of gene expression from nuclear-localized DNA. This chart is reproduced here:

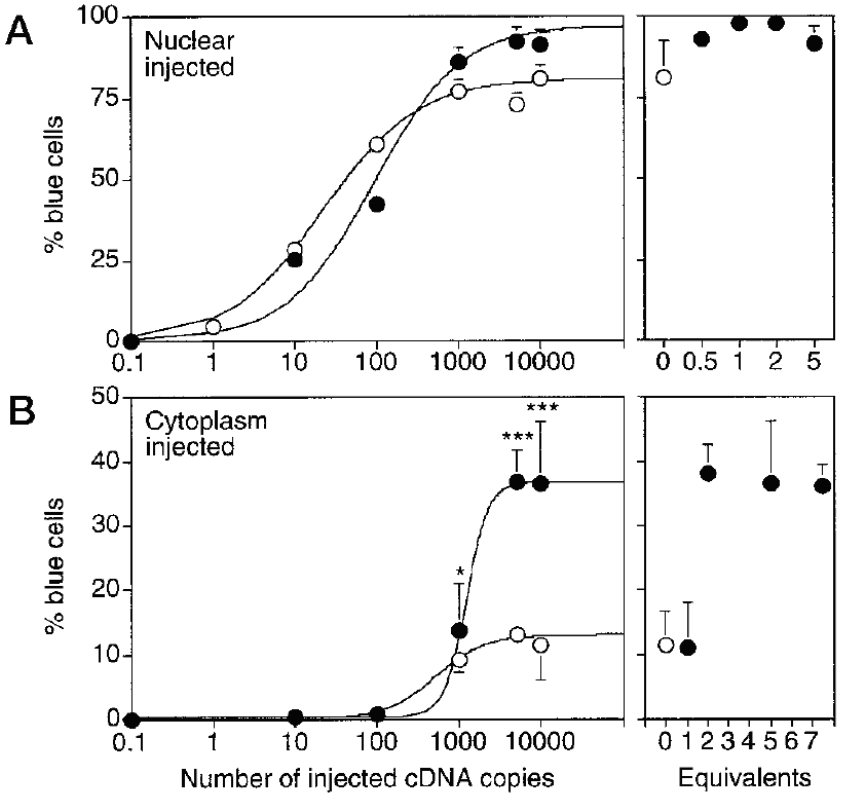


FIG. 2. **Efficiency of intranuclear (*A*) and cytoplasmic (*B*) cDNA microinjections in COS-7 cells.** In *A* and *B* (*left panels*) dose-effect relations with naked pCMV-LacZ cDNA (*open symbols*) or with PEI·cDNA complexes at 5 equivalents (*filled symbols*) are shown. A variable number of injected plasmid copies as indicated on the *x* axis was used. Data are mean ± S.E., with *n* between 178 and 650. *Right panels* show the effects of varying the PEI/cDNA ratio on injection efficiency. Cells were injected with 10,000 plasmid copies. *Stars* indicate significant difference, with $p < 0.05$ (*) or $p < 0.001$ (***) between cells injected with naked DNA and cells injected with DNA complexed with PEI.

The relevant part of the chart is panel A (open circles) that shows the percentage of protein-expressing cells as a function of a number of microinjected naked cDNA molecules. One can see that only a small fraction of cells expresses

the gene when one cDNA is injected, meaning that not all DNA that is localized to the nucleus is transcriptionally-active. Indeed, only cDNA numbers between 100-1000 molecules per nucleus result is high proportion of protein-expressing cells. This suggests that on average, one out of 100 DNA molecules already localized to the nucleus is transcriptionally-active.

Combining all the above evidence, we can estimate that on average out of 10,000 plasmids delivered to a cell using lipofection, one is transcriptionally active. Applying this factor to Fig. 8 from Wen-Chi Tseng *et al.*, we find that the number of active plasmids in highly-expressing cells is 3x10^6/10^5 = 30. Interestingly, we get roughly the same number of plasmids if we take the assumption of Schwake *et al.* that one plasmid can generate 10^6 GFP molecules per cell. In Fig. 8, the brightest cells have 25x10^6 GFP molecules, that is, 25 plasmids – a remarkable consistency between different pieces of evidence given their complete independence!

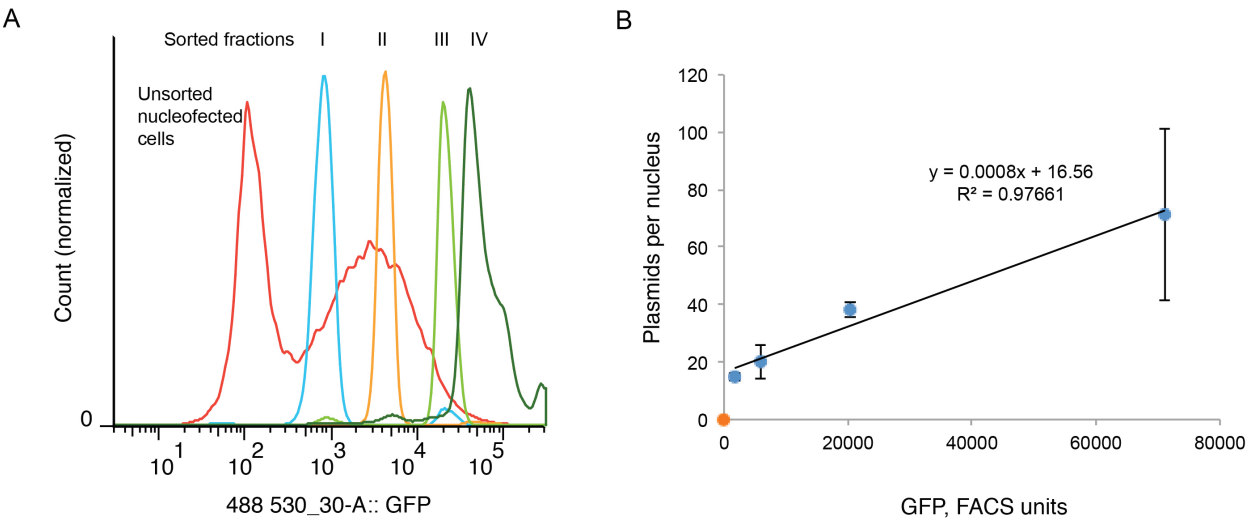
We stress here that we assume that DNA delivery yield (transcriptionally-active plasmids/total number of plasmids) is constant across the entire range of plasmid copies.

B. Evidence based on new experiments.

While partially satisfied with our analysis of published work, we decided to corroborate these conclusions with our own experiments. We went in the footsteps of *Richard N. Cohen et al.* (**paper 2** above) and decided to fractionate our transfected cell population based on fluorescent intensity and perform quantitative measurement of the nuclear-localized plasmids. Similar experiment was done in that report, but the numbers of plasmids were in the order of 1000s, even in low-expressing cells. We considered this to be a possible artifact of lipofection and the fact that not all nuclear DNA is transcriptionally-active. Therefore we decided to use a more direct DNA delivery method: nucleofection, which is a variant of electroporation and that delivers naked DNA to cells.

The experiment is fully-described in the manuscript and the supplementary information. Briefly, Tet-On HEK 293 cells were nucleofected with the transcriptional circuit and sorted into four fractions based on their ZsGReen levels. Nuclei were isolated from those cells and the extracted DNA was analyzed for the transfected plasmids as well as for an internal control gene using qPCR. We then plotted the normalized plasmid number as a function of mean ZsGreen level in those cells.

The figure (Supp. Fig. 7) is reproduced here:



Panel A shows post-FACS flow cytometry analysis of the sorted cells overlaid on the original unsorted population, and panel B shows the results of qPCR measurements, normalized to an internal control gene. We find a striking qualitative similarity to the shape of response in Fig. 8 from Wen-Chi Tseng et al, namely an apparent threshold effect. We attribute this to a fixed number of transcriptionally-inactive plasmids present in each cell, perhaps complexed to the membrane or simply partially degraded (yet giving a qPCR signal). In other words, we can safely assume that

Active plasmids = Total number of plasmids – constant

And therefore

Active plasmids = Constant x GFP.

Again, the results of our experiment are consistent with the analysis of the published literature. In our hands the number of active plasmids in the brightest cells is about 55, same order of magnitude as deduced above from published works. Therefore, the choice of 100 plasmids as the range in our simulations is reasonable.

Comment 2.1ii

1(ii). On that note, Schwake et al used a similar transfection protocol (Lipofectamin 2000 and similar amount of plasmid). Their conclusion was that the number of active vector complexes (vectors which express a protein) is on the order of one. Can the authors determine the level of active complexes in their system? This is important since the authors assume a wide distribution for the plasmid copy number and all subsequent interpretation of the results as well as the basis for the simulations and noise analysis relies on this assumption. Also, if this is indeed the case, then a significant part of the noise will result from the small number of plasmids. The authors indeed find a significant noise which can not be explained based on their various control circuits.

Response 2.1ii

We believe that our Response 2.1i (see above) addresses this question. Briefly, the number of active complexes or active plasmids in our and other systems ranges from one to 30-50. Schwake et al in their paper estimated the range to be between one and 15 plasmids. The discrepancy in numbers could be due to different DNA construct used, different delivery methods or different promoter strength. We agree that part of the noise can result from low plasmid numbers, even though this should have been accounted for in our control circuits.

Comments 2.2i and 2.2ii

2(i). FACS data.

In fig 3A, the authors show that with full functional repression by LacI (WT), gene expression of the output saturates quickly with respect to the plasmid copy number (input). The adaptive feature gradually decreases when decreasing LacI binding (the different mutants and the negative control). While the increase in the expression of DsRed is obvious when reducing LacI binding, it is not obvious why the maximal expression level of the input (ZsGreen) decreases as the maximal level of DsRed increases. Can the authors explain this observation? One possibility which the authors ought to rule out is that LacI binding to the mutated (and to the negative ctrl) sites of LacO somehow interferes with the expression of the divergent DsRed gene.

2(ii) In Fig 4, the authors show saturation for the post-transcriptional IFFLs. Again, the maximal expression levels for the negative control are significantly lower. In fact, as shown in fig 4c, there is a similar, almost perfect linear relationship, between the input and the output for the IFFL and the negative control for the same range of input levels. Thus, the negative control does not show that it does not adapt as opposed to the 'WT' circuit that reaches saturation. Can the authors explain that? Could it be that if the negative control circuits had higher expression levels a similar saturation will be observed? If so, it will suggest that IFFLs are not required for the adaptive feature in these synthetic circuits and that the adaptation is somehow inherent to this synthetic system.

Response 2.2

Indeed, we observe the difference in the input range between the fully-active circuits and the controls. To the first point about possible interference LacI with divergently-expressed LacI-IRES-ZsGreen due to binding to a mutated LacO sequence, we respectfully disagree with such interpretation. Indeed, the binding is strongest when the LacO sequence is not mutated. If there were any divergent effect on the expression of LacI-IRES-ZsGreen, it would be strongest in this context. Otherwise we have to assume that increasingly weaker binding induces increasingly strong effect on the divergent expression.

In addition to the above argument, the largest change in the input range is observed in the post-transcriptional IFFL. All the interactions between this circuit’s components are post-transcriptional, and the circuit generates three active components: a DsRed mRNA that is translated into DsRed protein (input); a

miR-FF3 that is spliced out of DsRed mRNA; and AmCyan mRNA that is targeted by miR-FF3 and is also translated to generate AmCyan reporter (output). Therefore it is highly unlikely that differences in the input range are due to decreased transcription rate of DsRed mRNA; instead, those differences must be caused at the translational level. Accordingly, we conjectured that simple reduction in protein translation from AmCyan mRNA due to RNAi (that in general can lead to mRNA degradation or inhibition of translation, or both), rather than any topology-specific circuit feature, leads to an increase in protein translation from the DsRed mRNA.

In other words, we made a reasonable assumption that in the post-transcriptional IFFL both the control and the circuit plasmids lead to the same DNA copy-number and newly synthesized mRNA abundance distribution. However circuit-induced change in steady-state output mRNA abundance leads to changes in protein translation rate from the input mRNA, meaning that the same number of plasmids and DsRed mRNA molecules will generate less DsRed protein in a control vector than in the circuit vector, decreasing the number of protein molecules per mRNA/plasmid in the control compared to the circuit. We speculate that this increase is caused by better availability of ribosomes that are otherwise translating AmCyan mRNA.

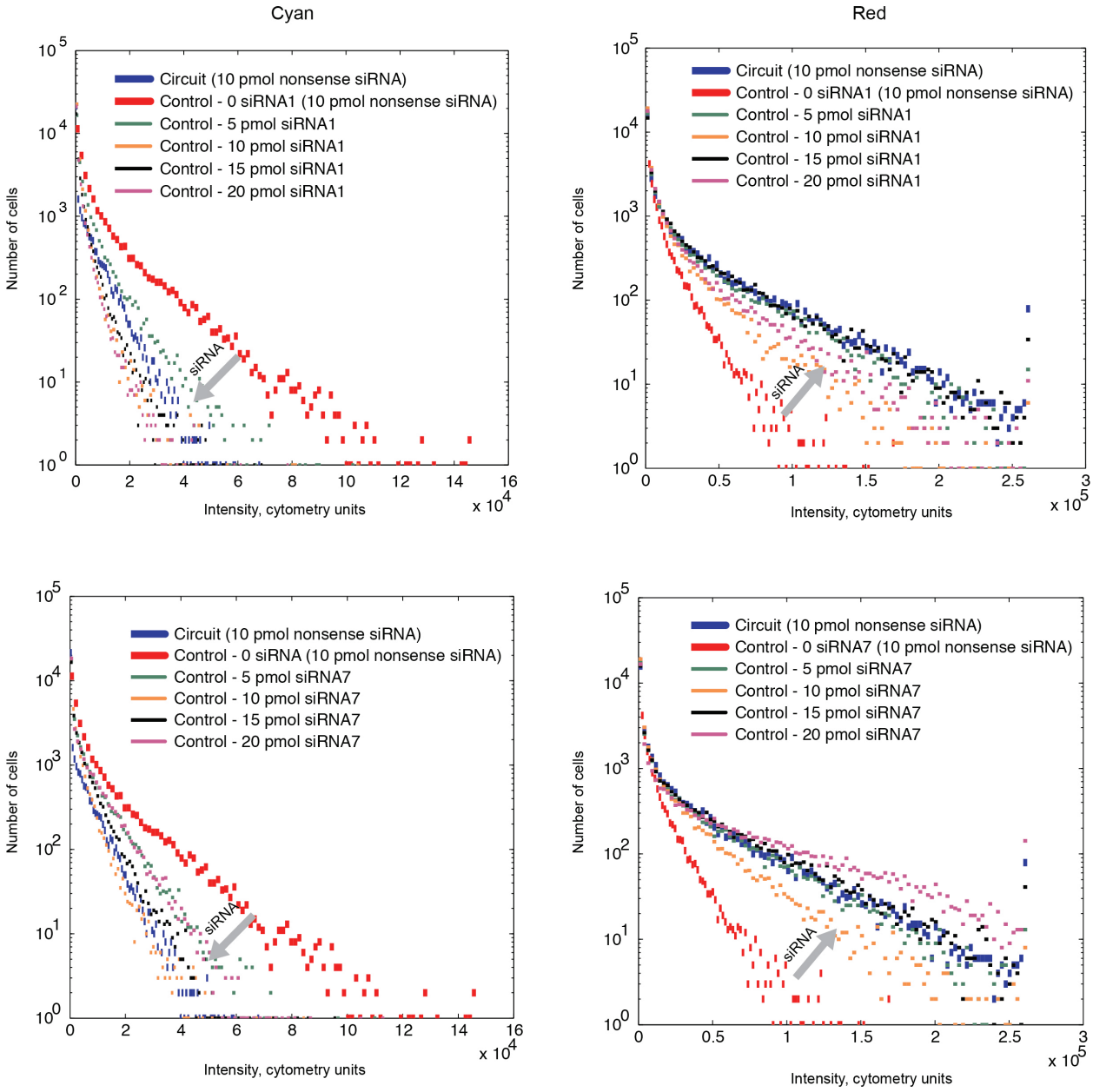
To test this assumption we conducted experiments using the posttranscriptional circuits. Because in the control circuit the input does not generate microRNA against the output, we targeted the AmCyan mRNA by siRNAs molecules complementary to AmCyan coding region to investigate the effect of generic AmCyan reduction on DsRed levels. The results are shown in the new Supplementary Figure 8, reproduced below. We show distributions of both DsRed and AmCyan markers in the circuit and the control (red and blue lines, respectively). The baseline distributions are consistent with previous observations, namely the Cyan is repressed in the circuit while DsRed (input) span is larger. We then incubated the control circuit with varying concentration of two different siRNAs that were found to be efficient Cyan downregulators in preliminary experiments. It can be seen that the Cyan is repressed to the levels observed in the circuit (green, orange, black and purple datasets) while at the same time DsRed distribution becomes similar to that observed in the circuit. We conclude that DsRed span increases in the circuit due to RNAi directed against the AmCyan output.

We emphasize that this observation does not affect our conclusion on the adaptive nature of the topology. Indeed, siRNA knock-down of AmCyan does not result in an adaptive behavior observed with the circuit, as shown in Supplementary Fig. 9 (we note that the siRNA experiments were performed with Lipofectamine 2000 due to poor siRNA transfection with Lipo LTX; Lipofectamine 2000 had been founds previously to be a sub-optimal transfection reagent for the post-transcriptional circuits).

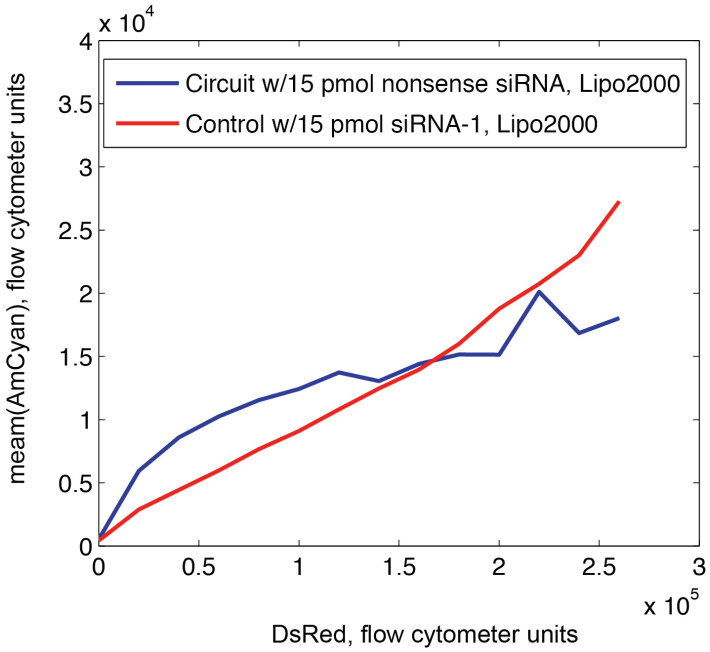
The same effect is likely to govern the change in the input range for the transcriptional circuits. In addition, we note that in these circuit the difference between the circuits and the controls is much less pronounced than in the post-transcriptional circuits and some of the apparent difference in the charts is due to data truncation.

We conclude:

(1) that the change in the input range is due to change in mRNA abundance of the output, either due to posttranscriptional repression by miRNA or transcriptional repression via LacI, effectively reducing the conversion coefficient between the input fluorescence and the plasmid copy number (plasmids per unit fluorescence) in the circuits and increasing it in the controls
(2) that this change cannot be the cause of the observed adaptive behavior in the circuits.



Supplementary Figure 8



Supplementary Figure 9

Comment 2.2iii

2(iii) Also, I could not find any referral to fig 4e in the text. There is also no description of what the two right most images of fig 4e represent (IFFL type I or II ?).

Response 2.2iii

We added the following sentence: “We also present in Figure 4E representative microscopy images of the signals dsRed and amCyan for the post-transcriptional feed-forward loops version I, II, and the negative control.”

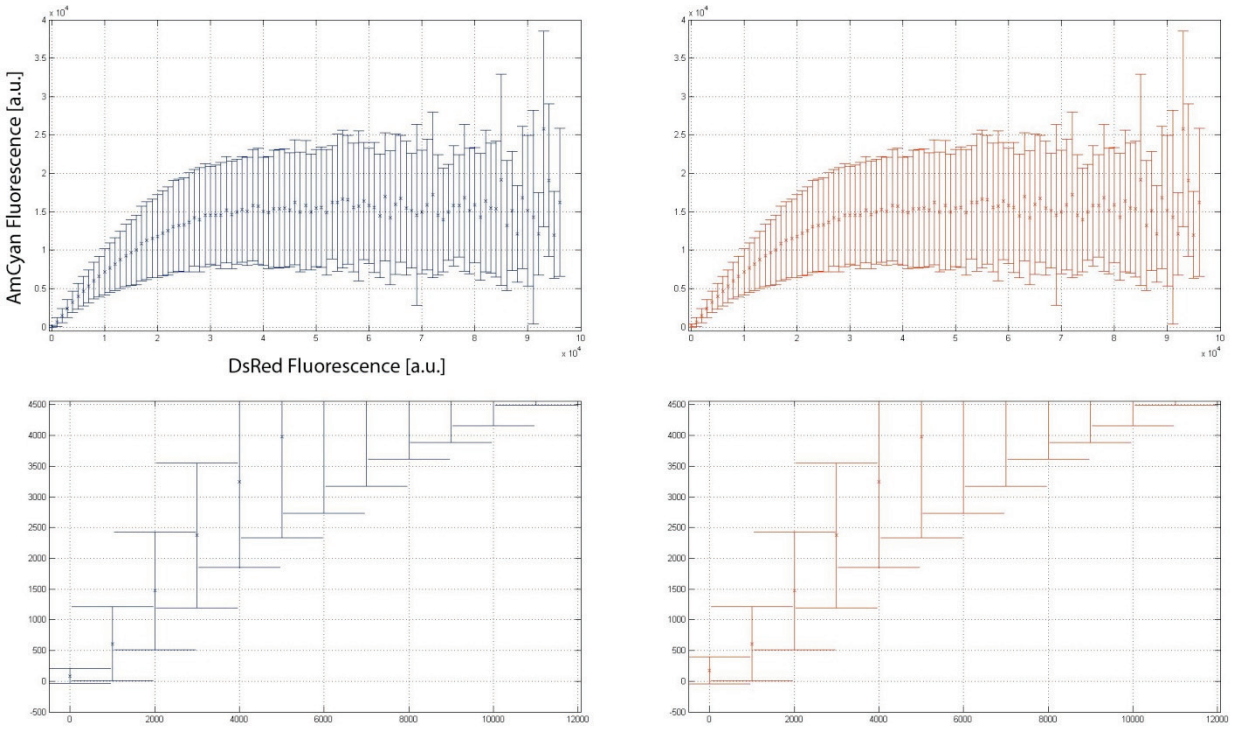
Comment 2.3

FACS analysis

In transformations only a small fraction of the cells eventually obtain the plasmid. Were the non-transformed cells excluded from the analysis? Can the authors show the background level of the cells (non-transformed) for the different channels for the same instrument setup? The authors restricted the analysis to counts which were less than half of the instruments' dynamic range to avoid biasing towards saturated data points, but there is no reference regarding the lower range of the dataset. For example, in figure S2A, it seems that a significant fraction of the cells have considerable DsRed levels (up to 2x10^4 a.u.) but the AmCyan levels are close to zero. Could it be that the lower range of the input-output linear dependence actually falls in the background autofluorescence of the cells?

Response 2.3

We did not restrict the data points while binning: for each interval of input values as defined by the individual bins, the output values of all cells whose input falls within this interval were collected and analyzed. The only restriction was in the very first bin, where we analyzed cells whose input level was above an auto-fluorescence of the negative control (which is very low compared to the entire range of the input). When we recalculated the response curves without any filtering out of cells, the results look virtually the same with a very small difference observed in the first bin:



Top and bottom left: The original data with truncated input range on top, and bottom the zoom in area of interest. Top and bottom right: The unprocessed data on top, and bottom the area of interest.

Minor:

Comment 2.4

Some of the statements are not clear or don't make sense:

(1) Page 6 (in the middle), the sentence " While strictly speaking...differences in reporter levels" does not sound logic after " synthesis and degradation rates,

". Should either add however or end the sentence and start a new one " Differnces in the copy-number...".

Response 2.4

The sentence was re-written as follows: “While strictly speaking, this reporter level also depends on many other potentially fluctuating parameters such as global synthesis and degradation rates, it is the differences in the copy-number that are the major source of cell-to-cell variability in transient transfections. Therefore as a first approximation the copy-number is considered as the sole contributor to the differences in reporter levels.“

Comment 2.5

(2) The last sentence starting on page 11 and continues on page 12 does not read well once reading ", are slightly sensitive...."

Response 2.5

We corrected the sentence as follows:

“The averaged input-output function of the tI1-FFL and the noise levels exhibit the following characteristics: (1) they are insensitive to changes in plasmid amounts used in transfections as well as Dox levels (Supplementary Figure S2E-K); (2) that are slightly sensitive to changes in transfection reagent (Supplementary Figure S2L and M); and (3) they do not change between 48 and 72 hours post-transfection (Supplementary Figure S2N and O).”

Comment 2.6

(3) page 13, typo in the 1st paragraph, "outut"

Response 2.6

Corrected, thank you.

Comment 2.7

(4) in SI, page 40 typo "explanded"

Response 2.7

Corrected, thank you.

Reviewer #3 (Remarks to the Author):

The authors analyze synthetic biological circuits in mammalian cells to explore levels of adaptation and fluctuations in output resulting from changes in DNA copy number. They focus on two types of post-transcriptional incoherent feed-forward loops and compare their behavior with a transcriptional FFL and a negative feedback loop. They use numerical and analytical modeling to generate predictions about the behavior of these various circuits. They predicted that all forms of FFLs would show adaptation with respect to DNA copy number variation, but as the inhibition is weakened the input-output behavior would become more linear. The analytical treatment provided additional constraints for when these results would hold true. They tested and validated the predictions experimentally using transient transfection of genetic constructs and flow cytometry. They show that the post-transcriptional IFFLs are less noisy and provide a greater range of adaptability compared with the transcriptional IFFL.

This work presents a clear analysis of a biologically important genetic motif. The manuscript shows a nice, complementary combination of computational, analytical, and experimental techniques to tackle an intriguing question.

Thank you!

Major comments:

Comment 3.1

1) The authors rely on a fluorescent read-out of the input to the circuit, which is the DNA copy number. However, there is no attempt at calibration between the two quantities. The only mention of this is during the discussion of model in which it is stated, "The amplitude of the measured fluorescence was arbitrarily assigned to correspond to 100 plasmid copies." As far as I know there is no support for such a connection in the literature. Providing some form of calibration would greatly strengthen the manuscript. One possibility would be to use a targeted stable knock-in of the genetic circuit into cells for calibration purposes.

Response 3.1

Essentially the same point was raised by reviewer 2; please see our comprehensive **response 2.2** above that included analysis of published work with new experiments, based on cell sorting and qPCR of transiently-transfected cells. Our conclusions support the original assertion regarding the linear relationship between the copy number and the fluorescence levels.

Comment 3.2

2) The analysis of the "biphasic" regime of the IFFLs is confusing. It seems that for conditions in which such biphasic behavior arises, a higher copy number results in a breakdown of adaptation. However, the experimental analysis of the

circuits shows the exact opposite - breakdown of adaptability occurs at low copy number. The authors should address these contradictory results.

Response 3.2

We believe that our experimental conditions (transfection efficiency) and circuit parameters are the reason we don't see the theoretically possible biphasic response. The model suggests that the biphasic behavior is a result of the slow degradation of dimers and tetramers as compared to the degradation of monomers. The saturation behavior, instead, is expected when degradation rates of all the species are similar and slow (i.e., on the order of by cell division rate) which is closer to the wild type case.

Comment 3.3

3) The results are encouraging for synthetic biologists. However, the conclusion that adaptation occurs only in the larger input regimes reduces the potential impact in regard to gene dosage compensation in natural systems, in which gene copy numbers are very small. Are there other systems which this study can shed light on their behaviors? i.e., systems with greater genetic copy numbers? A discussion of such systems will increase the impact of the manuscript.

Response 3.3

We agree with the reviewer on this point. We added the following paragraph in the discussion:

"While qualitatively our circuits exhibit dose compensation and adaptation for increased DNA amount, the actual input values that exhibit efficient compensation are relatively high. In endogenous systems compensation may be required when the copy number increases from two to three or two to four, or even one to two, and such compensation is unlikely to be caused by the exact replicas of our circuits. Having said that, there are examples in cancer where the amplification is 8-fold, that is, 16 copies of the gene are present in cells (Keyomarsi & Pardee, 1993), well within adaptation range in our circuits. If such dramatic amplifications are biologically feasible, the cell might pre-empt their negative effects by evolving a repressor or microRNA binding site targeted by closely-located and potentially co-amplified negative regulators. In addition, future research will address the fine-tuning of the input-output response in our circuits and attempt to uncover the "knobs" that control the half-saturation point, or EC_{50}, of the input. It is not infeasible that proper adjustments to circuit architecture could lower this value significantly in a rational manner and cause adaptive behavior even at very low copy numbers of one to four."

Minor comments:

Comment 3.4

1) It would be helpful to include diagrams and more detailed descriptions of the negative control circuits in the main text, instead of just the Supplementary

Materials. It was especially difficult to track down the negative control for the post-transcriptional IFFLs in Figure 4A and B.

Response 3.4

We added the circuit diagrams for the control plasmids in Fig. 1 and modified the legend accordingly.

Comment 3.5

2) Why are there different ranges for the inputs in the different experiments in Fig. 3A? One would expect that under the same transfection conditions, expression of the green fluorophore should have a comparable range, since the mutation in the LacO should only affect DsRed expression.

Response 3.5

Please see our **Response 2.2** above to Reviewer 2 Comment 2.2 i and 2ii. Summarizing, the apparent reduction of the input range is partially due to the truncation of the flow cytometry data (we restricted our analysis to those bins whose mean output value is less than half the dynamic range of the instrument (125K units) in order to avoid biased calculations of mean and standard deviation by including saturated data points). Indeed, the entire input range is only mildly reduced with decreased LacI activity. However, we do observe substantial reduction in the input range in the post-transcriptional circuits. We performed additional experiments to address this phenomenon, detailed in **Response 2.2**. Our conclusion is that what we observe is an increase in protein translation from the input-encoding mRNA in the circuit as a result of decrease in protein translation from the output-encoding mRNA. This in itself might be caused by a limiting number of available ribosomes of initiation factors for very high plasmid copy numbers at the extreme end of the input range.

Comment 3.6

3) I believe the caption of Supp Fig S2N should refer to red and black, rather than green and brown.

Response 3.6

Corrected, thank you.

References cited in this response

Keyomarsi K, Pardee AB (1993) Redundant cyclin overexpression and gene amplification in breast-cancer cells. *Proceedings of the National Academy of Sciences of the United States of America* **90:** 1112-1116

Nakamoto M, Jin P, O'Donnell WT, Warren ST (2005) Physiological identification of human transcripts translationally regulated by a specific microRNA. *Human Molecular Genetics* **14:** 3813-3821

Stenberg P, Lundberg LE, Johansson AM, Ryden P, Svensson MJ, Larsson J (2009) Buffering of Segmental and Chromosomal Aneuploidies in Drosophila melanogaster. *Plos Genetics* **5:** 10

Tsang J, Zhu J, van Oudenaarden A (2007) MicroRNA-mediated feedback and feedforward loops are recurrent network motifs in mammals. *Mol Cell* **26:** 753-767

Veitia RA, Bottani S, Birchler JA (2008) Cellular reactions to gene dosage imbalance: genomic, transcriptomic and proteomic effects. *Trends in Genetics* **24:** 390-397

Acceptance letter 06 June 2011

Thank you again for sending us your revised manuscript. We are now satisfied with the modifications made and I am pleased to inform you that your paper has been accepted for publication.

SOURCE DATA In view of the quantitative nature of your study, we would strongly encourage you to forward us by email a zipped folder with 'source data files' for figure panels representing quantitative experimental data (ie panels in Figure 3, 4 and 5). Source data are directly linked to specific figure panels so that interested readers can directly download the associated 'source data' (see, for example, <http://tinyurl.com/365zpej>), for the purpose of alternative visualization, re-analysis or integration with other data. I paste below the respective section of you instruction to authors and general formating guidelines are available at <http://www.nature.com/msb/authors/source-data.pdf>.

Thank you very much for submitting your work to Molecular Systems Biology.

Sincerely,

Editor
Molecular Systems Biology

From the instruction to authors (http://www.nature.com/msb/authors):

Source data for figures:

Molecular Systems Biology strongly encourages authors to upload the 'source data'-for example, tables of individual numerical values and measurements-that were used to generate figures. These files are separate from the traditional supplementary information files and are submitted using the "figure source data" option in the tracking system. Source data are directly linked to specific figure panels so that interested readers can directly download the associated 'source data' (see, for example, <http://tinyurl.com/365zpej>), for the purpose of alternative visualization, re-analysis or integration with other data.

The data should at usable by readers who understand the topic of the paper and should, at minimum, allow users to redraw the figures using an alternative graphical representation. Data should be annotated to provide essential information values included, providing a clear link to the figure legends in the main paper. Additional formatting guidelines for 'source data' are available for download <http://www.nature.com/msb/authors/source-data.pdf>.

Reviewer #2 (Remarks to the Author):

The authors addressed all the concerns. I therefore recommend the publication of the manuscript.

Reviewer #3 (Remarks to the Author):

The current study has been significantly improved since the original submission. All the concerns I had while reviewing this study have been satisfactory addressed. The addition of proper controls and analysis make the conclusions stronger and convincing. In cases where I, or the other reviewers, found the data or analysis confusing (for example, my previous comment 3.2) it would be helpful to clarify the results also in the paper and not only in the "response to reviewers", since other might be confused as well.

